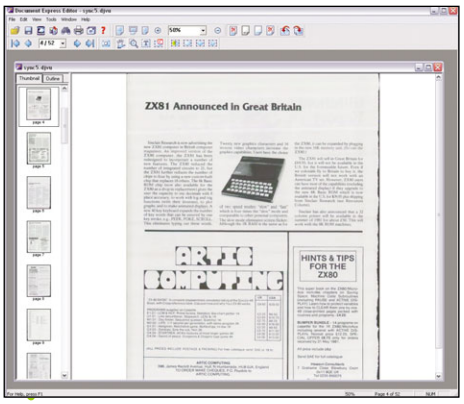
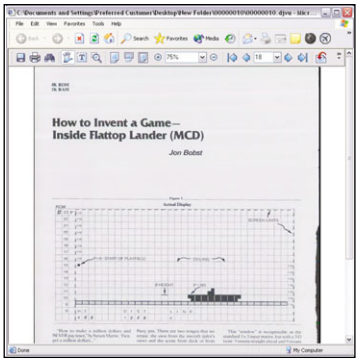
# LizardTech Document Express 5.0 With DjVu Professional Edition




**Document Express 5.0 With DjVu Professional Edition**
$295
LizardTech
www.lizardtech.com




Document Express 5.0 hits and misses with its scans.

Electronic document management and archiving is a hot topic, not only for businesses but also for public online archives. How can you effectively digitize and distribute piles of paper documents such as years worth of *CPU* magazines? Document Express 5.0, an application for creating digital archives using the DjVu file format, is one answer.

DjVu is a file format used for the delivery of scanned and digital documents. Most people are familiar with PDFs for distributing electronic documents, and DjVu is similar. PDFs do their best work with electronically produced documents. DjVu was designed specifically to store scanned documents. It can produce files that are significantly smaller than PDFs, JPEGs, PNGs, and other files created with common formats.

Like PDF, DjVu is an open format that's controlled by one company. (LizardTech bought the rights to DjVu from AT&T Labs in 2000.) Other companies can create software for viewing and creating DjVu files; a handful of open-source creators and viewers are available. DjVu isn't as widely used as PDF, but many well-known Web sites (the Internet Archive and NewspaperArchive.com) do use it.

Document Express 5.0 is the app used to create DjVu files. The Professional Edition is intended for small offices that need to scan documents with a TWAIN scanner or create digital documents by printing from any program. (For heavy output, a $6,495 Enterprise Edition is also available.)

I tested the app with a flatbed scanner and an automatic document feeder. The program worked with both but forced me to enter the flatbed scanning utility again and again for each page I wanted to scan. Once your document is scanned, you can reorder pages, highlight sections, add annotations, and create hyperlinks.

Although the program doesn't make it obvious, it can also create DjVu files from image files that you already have. I fed it 50 JPEG files that had been previously scanned with a high-speed drum scanner. The conversion wasn't particularly fast, taking 25 minutes for 50 pages. The resulting file was 6.4MB, a fraction of the original JPEGs, which weighed in at 35MB. DjVu is a lossy format, but in all my tests, the program produced perfectly readable color images.

When viewing a DjVu document, you see the original page image. The program's OCR support provides for the text to be encoded within the file. As a result, you can search for text within a graphical document, which is incredibly handy for finding data buried in a lengthy document. The program's OCR isn't perfect, but because you don't usually see the text directly, it really doesn't need to be. In addition to scanning, you can create DjVu documents from any app that can print using the virtual printer driver.

In order to view DjVu documents, users need either a standalone reader (which is available for Windows 95/98/NT/2000/XP) or a browser plug-in. Both are free to distribute to other users. As a rule, browser plug-ins are a pain, and many Web users simply avoid using sites that rely on them. This is perhaps the biggest strike against the DjVu technology. (To be fair, PDF files also require a plug-in for most browsers, although it is common enough to come preinstalled with some.) There are ways around the plug-in requirement, including a Java-based viewer and a server add-on, both of which are added expenses.

Plug-ins are available for a variety of platforms, including Windows running Internet Explorer, Mac OS X running Safari, Mac OS 9, and many versions of Linux/Unix. The third-party, open-source Linux/Unix plug-in (available from djvulibre.djvuzone.org) works with Netscape, Mozilla, Galeon, Konqueror, and Opera.

Document Express doesn't require much horsepower; the recommended system requirements are a 400MHz processor with 128MB of RAM, so you can easily put an old PC to work on converting documents. ▲

**by Kevin Savetz**

CPU Ranking: 0 = Absolutely Worthless 2.5 = Absolutely Average 5 = Absolutely Perfect